CLAIMS

1. A method for determining where to store object replicas, the method comprising:

receiving an indication of a homeless replica of an object, wherein the object has a plurality of replicas including the homeless replica;

determining an initial placement for the homeless replica, wherein the initial placement is one of a plurality of devices in a system;

evaluating whether any replicas of the object can be swapped with one of a plurality of replicas of another object and not reduce a combined object availability of the two objects; and

swapping a replica of the object with the one of the plurality of replicas of the other object only if the swapping does not reduce the combined object availability of the two objects.

2. A method as recited in claim 1, wherein the replica of the object is stored on a first computing device, wherein the one of the plurality of replicas of the other object is stored on a second computing device, and wherein swapping the replicas comprises moving the replica of the object to the second computing device and moving the one of the plurality of replicas of the other object to the first computing device.

3. A method as recited in claim 1, wherein the method is implemented by multiple computing devices in a serverless distributed file system.

4.   A method as recited in claim 1, wherein both the object and the other object have a same amount of replicas.

5.   A method as recited in claim 1, wherein the object comprises a file.

6.   A method as recited in claim 1, wherein the object comprises a plurality of files.

7.   A method as recited in claim 1, wherein the object comprises a portion of a file.

8.   A method as recited in claim 1, wherein the object comprises a directory.

9.   A method as recited in claim 1, wherein the determining comprises randomly selecting the initial placement for the homeless replica.

10.   A method as recited in claim 1, wherein the evaluating comprises:

       determining, for each possible swapping of replicas of the objects, an object availability for each object; and

       checking whether any of the possible swappings result in a combined object availability of the two objects that is greater than the combined object availability prior to the evaluating.

11. A method as recited in claim 1, wherein combined object availability is not reduced if the availability of the object and the availability of the other object are no further from one another than they were prior to the evaluating.

12. A method as recited in claim 1, wherein the swapping results in increasing an availability of the object and decreasing an availability of the other object.

13. A method as recited in claim 1, wherein the swapping results in decreasing an availability of the object and increasing an availability of the other object.

14. A method as recited in claim 1, implemented by a computing device and further comprising selecting, prior to the evaluating, as the object one of a plurality of objects that the computing device is responsible for maintaining.

15. A method as recited in claim 14, wherein the selecting comprises randomly selecting one of the plurality of objects.

16. A method as recited in claim 14, wherein the selecting comprises selecting the one of the plurality of objects that has the lowest object availability.

17. A method as recited in claim 1, wherein the other object is one of a plurality of objects that a computing device is responsible for maintaining, and wherein the other object is randomly selected from the plurality of objects.

18. A method as recited in claim 1, wherein the other object is one of a plurality of objects that a computing device is responsible for maintaining, and wherein the other object is the one of the plurality of objects that has the highest object availability.

19. One or more computer readable media having stored thereon a plurality of instructions that, when executed by one or more processors of a computing device, causes the one or more processors to perform acts comprising:

working, in conjunction with one or more other processors of another computing device, to determine whether a replica of a file managed by the computing device and a replica of another file managed by the other computing device can be swapped with one another to bring an availability of the file and an availability of the other file closer; and

swapping the replica of the file and the replica of the other file only if the swapping brings the availability of the file and the availability of the other file closer.

20. One or more computer readable media as recited in claim 19, wherein the swapping comprises communicating with a first device on which the replica of the file is stored and a second device on which the replica of the other file is stored to have the first device transfer the replica of the file to the second device and delete the replica of the file on the first device, and to have the second device transfer the replica of the other file to the first device and delete the replica of the other file on the second device.

21. One or more computer readable media as recited in claim 19, wherein the one or more processors are part of a device in a serverless distributed file system.

22. One or more computer readable media as recited in claim 19, wherein the file represents a plurality of files.

23. One or more computer readable media as recited in claim 19, wherein the file represents a portion of another file.

24. One or more computer readable media as recited in claim 19, wherein the file represents a folder.

25. One or more computer readable media as recited in claim 19, wherein the plurality of instructions further cause the one or more processors to perform acts comprising:

receiving an indication of a homeless replica of the file; and

determining an initial placement, on one of a plurality of additional computing devices, for the homeless replica.

26. One or more computer readable media as recited in claim 25, wherein determining the initial placement comprises randomly selecting the initial placement for the homeless replica.

27. One or more computer readable media as recited in claim 25, wherein the plurality of additional computing devices exclude any computing device on which a replica of the file is already stored.

28. One or more computer readable media as recited in claim 19, wherein the file is selected randomly from a plurality of files managed by the computing device, and wherein the other file is selected randomly from another plurality of files managed by the other computing device.

29. One or more computer readable media as recited in claim 19, wherein the file is selected as the one of a plurality of files managed by the computing device having a lowest file availability, and the other file is selected randomly from another plurality of files managed by the other computing device.

30. One or more computer readable media as recited in claim 19, wherein the file is selected as the one of a plurality of files managed by the computing device having a lowest file availability, and the other file is selected as the one of another plurality of files managed by the other computing device having a highest file availability.

31. One or more computer readable media as recited in claim 19, wherein the computing devices is part of a directory group that is collectively responsible for managing a plurality of files including the file, and wherein the other computing device is part of another directory group that is collectively responsible for managing another plurality of files including the other file.

32. One or more computer readable media as recited in claim 19, wherein the one or more processors of a computing device and the one or more other processors of another computing device are the same processors of the same computing device.

33. A serverless distributed file system comprising:

a first plurality of computing devices storing files;

a second plurality of computing devices managing storage of the files;

wherein a first computing device of the second plurality of computing devices selects a file for which it manages storage and communicates with a second computing device of the second plurality of computing devices to determine whether a replica of the file and a replica of another file for which the

second computing device manages storage can be swapped in order to improve a combined file availability; and

if the replicas can be swapped to improve the combined file availability, then instructing the one of the first plurality of computing devices on which the replica of the file is stored to transfer the replica of the file to the one of the first plurality of computing devices on which the replica of the other file is stored, and instructing the one of the first plurality of computing devices on which the replica of the other file is stored to transfer the replica of the other file to the one of the first plurality of computing devices on which the replica of the file is stored.

34. A serverless distributed file system as recited in claim 33, wherein the second plurality of computing devices comprise one or more directory groups.

35. A serverless distributed file system as recited in claim 33, wherein the second plurality of computing devices further receives an indication of a homeless replica of the file, and randomly selects, as a computing device on which to store the homeless replica, one of the first plurality of computing devices on which no other replica of the file is already stored.

36. One or more computer readable media having stored thereon a plurality of instructions that, when executed by one or more processors of a computing device, causes the one or more processors to:

initially place replicas of a file on different ones of a plurality of devices using a first process; and

subsequently improve the placement of replicas of a plurality of files by evaluating whether any replicas of a first file can be swapped with any replicas of a second file without a reduction in the combined file availability of the first and second files, and swapping a replica of the first file with a replica of the second file if the swapping results in no reduction in the combined file availability of the first and second files.

37. One or more computer readable media as recited in claim 36, wherein the first process comprises random selection.

38. One or more computer readable media as recited in claim 36, wherein swapping the replica of the first file with the replica of the second file comprises communicating with a first device on which the replica of the first file is stored and a second device on which the replica of the second file is stored to have the first device transfer the replica of the first file to the second device and delete the replica of the first file on the first device, and to have the second device transfer the replica of the second file to the first device and delete the replica of the second file on the second device.

39. One or more computer readable media as recited in claim 36, wherein the first file is selected randomly from a first plurality of files, and wherein the second file is selected randomly from a second plurality of files.

40. One or more computer readable media as recited in claim 36, wherein the first file is selected as the one of a first plurality of files having a lowest file availability, and the second file is selected randomly from a second plurality of files.

41. One or more computer readable media as recited in claim 36, wherein the first file is selected as the one of a first plurality of files managed having a lowest file availability, and the second file is selected as the one of a second plurality of files having a highest file availability.

42. A method, implemented in a directory group, the method comprising:

selecting another directory group to participate with in a replica placement improvement process;

selecting a file maintained by the directory group;

determining whether exchanging a replica of the file with a replica of another file maintained by the other directory group will increase a combined file availability of the files; and

having the replica of the file and the replica of the other file exchanged if exchanging the replicas will increase the combined file availability of the files.

43. A method as recited in claim 42, further comprising:

receiving, at the directory group, an indication of a homeless replica of the file; and

selecting, as a computing device on which to store the homeless replica, one of a plurality of computing devices on which no other replica of the file is already stored.

44. A method as recited in claim 42, wherein the selecting comprises randomly selecting.

45. A method as recited in claim 42, wherein having the replica of the file and the replica of the other file exchanged comprises communicating with a first device on which the replica of the file is stored and a second device on which the replica of the other file is stored to have the first device transfer the replica of the file to the second device and delete the replica of the file on the first device, and to have the second device transfer the replica of the other file to the first device and delete the replica of the other file on the second device.

46. A method as recited in claim 42, wherein the file is selected randomly from a plurality of files managed by the directory group, and wherein the other file is selected randomly from another plurality of files managed by the other directory group.

47. A method as recited in claim 42, wherein the file is selected as the one of a plurality of files managed by the directory group having a lowest file availability, and the other file is selected randomly from another plurality of files managed by the other directory group.

48. A method as recited in claim 42, wherein the file is selected as the one of a plurality of files managed by the directory group having a lowest file availability, and the other file is selected as the one of another plurality of files managed by the other directory group having a highest file availability.